Claims

WHAT IS CLAIMED IS:

1. A method comprising:

selecting a modeling parameter from a plurality of modeling parameters characterizing a mixture of Student distribution components;

computing a tractable approximation of a posterior distribution for the selected modeling parameter based on an input set of data and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

computing a lower bound of a log marginal likelihood as a function of current estimates of the posterior distributions of the modeling parameters, the current estimates of the posterior distributions of the modeling parameters including the computed tractable approximation of the posterior distribution of the selected modeling parameter; and

generating a probability density modeling the input set of data, the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters, if the lower bound is satisfactorily optimized.

2. The method of claim 1 wherein the computing operations comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing a

tractable approximation and computing a lower bound using the newly selected modeling parameter, if the lower bound is not satisfactorily optimized in the first iteration.

3. The method of claim 1 wherein computing a lower bound comprises:

computing the lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters.

4. The method of claim 1 wherein computing a tractable approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution of the selected modeling parameter.

5. The method of claim 1 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

6. The method of claim 1 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

7. The method of claim 1 wherein one of the plurality of modeling parameters represents a labeling parameter of the Student distribution components.

8. The method of claim 1 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

9. The method of claim 1 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

10. The method of claim 1 wherein generating a probability density comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the number of degrees of freedom of each Student distribution component.

11. The method of claim 1 further comprising:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

12. The method of claim 1 wherein the input set of data represents auditory speech data from an unknown number of speakers, and further comprising determining a correct number of speakers from the probability density modeling the input set of data.

13. The method of claim 1 wherein the input set of data represents image segmentation data from images having regions of different characteristics.

14. A computer program product encoding a computer program for executing on a computer system a computer process, the computer process comprising:

selecting a modeling parameter from a plurality of modeling parameters characterizing a mixture of Student distribution components;

computing a tractable approximation of a posterior distribution for the selected modeling parameter based on an input set of data and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

computing a lower bound of a log marginal likelihood as a function of current estimates of the posterior distributions of the modeling parameters, the current estimates of the posterior distributions of the modeling parameters including the computed tractable approximation of the posterior distribution of the selected modeling parameter; and

generating a probability density modeling the input set of data, the probability density including the mixture of Student distribution components , the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters, if the lower bound is satisfactorily optimized.

15. The computer program product of claim 14 wherein the computing operations comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing a tractable approximation and computing a lower bound using the newly selected

modeling parameter, if the lower bound is not satisfactorily optimized in the first iteration.

16. The computer program product of claim 14 wherein computing a lower bound comprises:

computing the lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters.

17. The computer program product of claim 14 wherein computing a tractable approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution of the selected modeling parameter.

18. The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

19. The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

20. The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a labeling parameter of the Student distribution components.

21. The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

22. The computer program product of claim 14 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

23. The computer program product of claim 14 wherein generating a probability density comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

24. The computer program product of claim 14 wherein the computer process further comprises:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

25. The computer program product of claim 14 wherein the input set of data represents auditory speech data from an unknown number of speakers, and further comprising determining a correct number of speakers from the probability density modeling the input set of data.

26. The computer program product of claim 14 wherein the input set of data represents image segmentation data from images having regions of different characteristics.

27. A system comprising:

a modeling parameter selector selecting a modeling parameter from a plurality of modeling parameters characterizing a mixture of Student distribution components;

a tractable approximation module computing a tractable approximation of a posterior distribution for the selected modeling parameter based on an input set of data and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

a lower bound optimizer module computing a lower bound of a log marginal likelihood as a function of current estimates of the posterior distributions of the modeling parameters, the current estimates of the posterior distributions of the modeling parameters including the computed tractable approximation of the posterior distribution of the selected modeling parameter; and

a data model generator generating a probability density modeling the input set of data, the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters, if the lower bound is satisfactorily optimized.

28. The system of claim 27 wherein the lower bound optimizer computes the lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters.

29. The system of claim 27 wherein the tractable approximation module computes a variational approximation of the posterior distribution of the selected modeling parameter.

30. The system of claim 27 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

31. The system of claim 27 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

32. The system of claim 27 wherein one of the plurality of modeling parameters represents a labeling parameter of the Student distribution components.

33. The system of claim 27 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

34. The system of claim 27 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

35. The system of claim 27 wherein the data model generator generates the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

36. The system of claim 27 further comprising:

a memory storing the current estimates of the posterior distributions of the modeling parameters.

37. The system of claim 27 wherein the input set of data represents auditory speech data from an unknown number of speakers, and further comprising determining a correct number of speakers from the probability density modeling the input set of data.

38. The system of claim 27 wherein the input set of data represents image segmentation data from images having regions of different characteristics.

39. A method comprising:

computing a tractable approximation of a posterior distribution for a selected modeling parameter of a plurality of modeling parameters characterizing a mixture of Student distribution components based on an input set of data and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized, the current estimates of the posterior distributions of the modeling parameters including the computed tractable approximation of the posterior distribution of the selected modeling parameter; and

modeling the input set of data by the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters.

40. The method of claim 39 wherein the computing operation and determining operation comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing a tractable approximation and computing a lower bound using the newly selected modeling parameter, if the lower bound is not satisfactorily optimized in the first iteration.

41. The method of claim 39 wherein the operation of determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized comprises:

computing a lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters and a variational posterior distribution; and

determining whether the lower bound satisfies a predetermined criterion of the selected modeling parameter.

42. The method of claim 39 wherein computing a tractable approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution.

43. The method of claim 39 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

44. The method of claim 39 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

45. The method of claim 39 wherein one of the plurality of modeling parameters represents a labeling parameter of the Student distribution components.

46. The method of claim 39 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

47. The method of claim 39 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

48. The method of claim 39 wherein modeling the input data comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

49. The method of claim 39 further comprising:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

50. A computer program product encoding a computer program for executing on a computer system a computer process, the computer process comprising:

computing a tractable approximation of a posterior distribution for a selected modeling parameter of a plurality of modeling parameters characterizing a mixture of Student distribution components based on an input set of data and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized, the current estimates of the posterior distributions of the modeling parameters including the computed tractable approximation of the posterior distribution of the selected modeling parameter; and

modeling the input set of data by the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters.

51. The computer program product of claim 50 wherein the computing operation and determining operation comprise a first iteration and further comprising:

selecting a different modeling parameter from the plurality of modeling parameters and repeating in a subsequent iteration the operations of computing a tractable approximation and computing a lower bound using the newly selected modeling parameter, if the lower bound is not satisfactorily optimized in the first iteration.

52. The computer program product of claim 50 wherein the operation of determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized comprises:

computing a lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters and a variational posterior distribution; and

determining whether the lower bound satisfies a predetermined criterion.

53. The computer program product of claim 50 wherein computing a tractable approximation of a posterior distribution comprises:

computing a variational approximation of the posterior distribution of the selected modeling parameter.

54. The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

55. The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

56. The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a labeling parameter of the Student distribution components.

57. The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

58. The computer program product of claim 50 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

59. The computer program product of claim 50 wherein modeling the input data comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

60. The computer program product of claim 50 wherein the computer process further comprises:

storing the current estimates of the posterior distributions of the modeling parameters in a storage location.

61. A system comprising:

a tractable approximation module computing a tractable approximation of a posterior distribution for a selected modeling parameter of a plurality of modeling parameters characterizing a mixture of Student distribution components based on an input set of data and a current estimate of a posterior distribution of at least one unselected modeling parameter in the plurality of modeling parameters;

an optimizer module determining whether current estimates of the posterior distributions of the modeling parameters are satisfactorily optimized, the current estimates of the posterior distributions of the modeling parameters including the computed tractable approximation of the posterior distribution of the selected modeling parameter; and

a data model generator modeling the input set of data by the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters.

62. The system of claim 61 wherein optimizer module computes a lower bound of the log marginal likelihood as a function of prior distributions of the modeling parameters and a variational posterior distribution, and determines whether the lower bound satisfies a predetermined criterion.

63. The system of claim 61 wherein the tractable approximation modules computes a variational approximation of the posterior distribution of the selected modeling parameter.

64. The system of claim 61 wherein one of the plurality of modeling parameters represents a mean of each of the Student distribution components.

65. The system of claim 61 wherein one of the plurality of modeling parameters represents a precision matrix of the Student distribution components.

66. The system of claim 61 wherein one of the plurality of modeling parameters represents a labeling parameter of the Student distribution components.

67. The system of claim 61 wherein one of the plurality of modeling parameters represents a scaling parameter of a precision matrix of the Student distribution components.

68. The system of claim 61 wherein one of the plurality of modeling parameters represents a mixing coefficients parameter of the Student distribution components.

69. The system of claim 61 wherein modeling the input data comprises:

generating the probability density including the mixture of Student distribution components, the mixture of Student distribution components being characterized by the current estimates of the posterior distributions of the modeling parameters and an estimate of the degrees of freedom of each Student distribution component.

70. The system of claim 61 further comprising:

a memory storing the current estimates of the posterior distributions of the modeling parameters.